Exploratory Analysis of Rain Fall Data in India for Agriculture

Team Id

PNT2022TMID10333

Faculty Mentor

Mrs. Anitha.P

Team Leader

Allenjoshuva .J (720719110114)

Team Member

Ranjith .R   (720719110074)

Venkatesan .S (720719110101)

Subash .D   (720719110090)

INTRODUCTION

Natural processes on Earth can be classified into several categories, including hydrological processes like storm waves and groundwater; biological processes like forest growth; atmospheric processes like thunderstorms and rainfall; human processes like urban development; and geological processes like earthquakes. The field of physical geography seeks to investigate the distribution of the different features/parameters that describe the landscape and functioning of the Earth by analysis in the processes that shape it. These features/parameters have been referred to as geophysical parameters in the literature. Rainfall is a key geophysical parameter that is essential for many applications in water resource management, especially in the agriculture sector. Predicting rainfall can help managers in various sectors to make decisions regarding a range of important activities such as crop planting, traffic control, the operation of sewer systems, and managing disasters like droughts and floods. In our country India, It depends upon the agriculture sector as a major contributor to the economy and as a source of food security. Hence, an accurate prediction of rainfall is needed to make better future decisions to help manage activities such as the ones mentioned above.

LITERATURE REVIEW

[1] E. A. Hussein et al. investigated identifying contributing factors and understanding the effect of these factors in causing the effectiveness of rainfall to have been one of the key focus areas of agriculture in research. The review covers several aspects which relate to the input into, output from, and methods used in the various systems devised in the literature for this purpose. The review specifically focuses on studies that use supervised learning for both regression and classification problems. This Literature was used to collect papers from 2016 to 2020, with the following key words like "machine learning" or "deep learning" and "precipitation prediction" or "rainfall prediction" or "precipitation nowcasting". Almost 1240 results were obtained, and of these only supervised rainfall prediction papers that used meteorological data from, e.g., radar, satellites, and stations were selected, while papers that used data from normal cameras, e.g., photographs were excluded. Hence, the conclusions and discussions of this chapter can be adapted to other parameters.

Algorithm used: PCA-ANN, PCA-MLR, ELM, DT, WT-SVM.

This algorithm uses the salient features in the data and various process in this method of prediction. Fast learning speed, Ease of implementation, minimal human intervention, Implementation efficiency.

Advantages: This prediction method used to forecast rainfall using ML models using five years data so more accurate prediction.

Disadvantages: However this study is limited in India. It is useful for short term prediction not suitable for long term prediction.

[2] Rainfall prediction is one of the challenging tasks in weather forecasting. Accurate and timely rainfall prediction can be very helpful to take effective security measures in advance regarding: ongoing construction projects, transportation activities, agricultural tasks, flight operations and flood situation, etc. Data m techniques can effectively predict the rainfall by extracting the hidden patterns among available features of past weather data. This research contributes by providing a critical analysis and review of latest data mining techniques, used for rainfall prediction. Published papers from year 2013-2017 from renowned online search libraries are considered for this research. This review will serve the researchers to analyse the latest work on rainfall prediction with the focus on data mining techniques and also will provide a baseline for future directions.

Algorithm used: Genetic based algorithm, Classification algorithm, MPL algorithm, SVR.

This algorithm manages to fit the best line within the threshold of values with the capability of generalization with high prediction accuracy. It can provide flexibility.

Advantages: This literature method uses Back Propagation Neural Network for prediction which is reported an acceptable accuracy. This will not affect analysis of data prediction.

Disadvantages: It modifies the technique that evaluates the results may not be as accurate as explained. This can be affects the analysis and results of this condition.

[3] The study of precipitation trends is critically important for a country like India whose food security and economy are dependent on the timely availability of water. In this work, monthly, seasonal and annual trends of rainfall have been studied using monthly data series of 135 years for 30 sub-regions in India. Half of the sub-divisions showed an increasing trend in annual rainfall, but for only three (Haryana, Punjab and Coastal Karnataka), this trend was statistically significant. Similarly, only one sub-region that indicated a significant decreasing trend out of the 15 sub-divisions showing decreasing trend in annual rainfall. In India, the monsoon months of June to September account for more than 80% of the annual rainfall. The majority of sub-divisions showed very little change in rainfall in non-monsoon months. For the whole of India, no significant trend was detected for annual, seasonal, or monthly rainfall. Annual and monsoon rainfall decreased, while pre-monsoon, post-monsoon and winter rainfall increased at the national scale.

Algorithm used: KNNs, SVM, ARMA, LR

No training for classification and regression. This algorithm has no explicit training step work during prediction. This method tries to find the best margin that separates the classes and it will reduce the risk of error of the data.

Advantages: The time series of rainfall data used in this study spans more than 100 years. Thus, the present analysis is significant improvement over the studies carried out previously shown.

Disadvantages: The five major regions of India showed not significant trend in annual, monthly rainfall it is difficult to predict in those areas. It takes some difficult process to predict the data analysis of rainfall.

[4] Rainfall is a key part of hydrological cycle and alteration of its pattern directly affect the water resources. The changing pattern of rainfall in consequence of climate change is now concerning issues to water resource manages and hydrologists. Srivastava et al. and Islam et al. are reported that the changes of rainfall quantities and frequencies directly changing the stream flow pattern and its demand, spatiotemporal allocation of run-of, ground water reserves and soil moisture. Consequently, these changes showed the widespread consequences on the water resource, environment, terrestrial ecosystem, ocean, bio-diversity, agricultural and food security. The drought and food like hazardous events can be occurred frequently because of the extreme changes of rainfall trend.

Algorithm used: Artificial Neural Network, Neural Network, Logistic Regression, Random Forest.

This is the group of algorithms that certify the underlying relationship in a set of data similar to the human brain. This helps to change the input so that network gives the best result without redesigning the output procedure.

Advantages: The writing was extricated with a thorough and careful process which shows the quality and culmination of this review anyway some significant applicable work may have been missed.

Disadvantages: Most of the incorporated and altered procedures were assessed by the authors themselves, so the genuine outcomes may not be as precise as clarified. This might influence the examination and results of this review.

REFERENCES

[1] E. A. Hussein et al. investigated identifying contributing factors and understanding the effect of these factors in causing the effectiveness of rainfall to

have been one of the key focus areas of agriculture in research,2016. Abbot, J., & Marohasy, J. (2017). Application of artificial neural networks to forecasting monthly rainfall one year in advance for locations within the Murray Darling basin, Australia. International Journal of Sustainable Development and Planning.

To link to this article: **Rainfall Prediction Using Machine Learning Models: Literature Survey (researchgate.net)**

[2] Rainfall prediction is one of the challenging tasks in weather forecasting. Accurate and timely rainfall prediction can be very helpful to take effective security measures in advance regarding. S. Zhang, L. Lu, J. Yu, and H. Zhou, "Short-term water level prediction using different artificial intelligent models," in 2016 5th International Conference on Agri-Geoinformatics,2016.

Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-67228-7.

To link to this article: https://doi.org/10.1080/02626667.2010.481373

[3] The study of precipitation trends is critically important for a country like India whose food security and economy are dependent on the timely availability of water. Aggarwal, P. K. (2007) Climate change: implications for Indian agriculture. Hydrology Review 22, 37–46. Roorkee: Indian National Committee on Hydrology. Arora, M., Goel, N. K. & Singh, P. (2005) Evaluation of temperature trends over India. Hydrology. Sci. J. 50(1), 81–93.

To link to this article: https://doi.org/10.1038/s41598-020-67228-7

[4] Rainfall is a key part of hydrological cycle and alteration of its pattern directly affect the water resources. Islam, T., Rico-Ramirez, M. A., Han, D. & Srivastava, P. K. A Joss–Waldvogel disdrometer derived rainfall estimation study by collocated tipping bucket and rapid response rain gauges. Atmospheric Science Letters 13, 139–150 (2012). 2. Gajbhiye, S., Meshram, C., Singh, S. K., Srivastava, P. K. & Islam, T. Precipitation trend analysis of Sindh River basin, India, from 102-year record (1901–2002). Atmospheric Science Letters 17(1), 71–77 (2015).